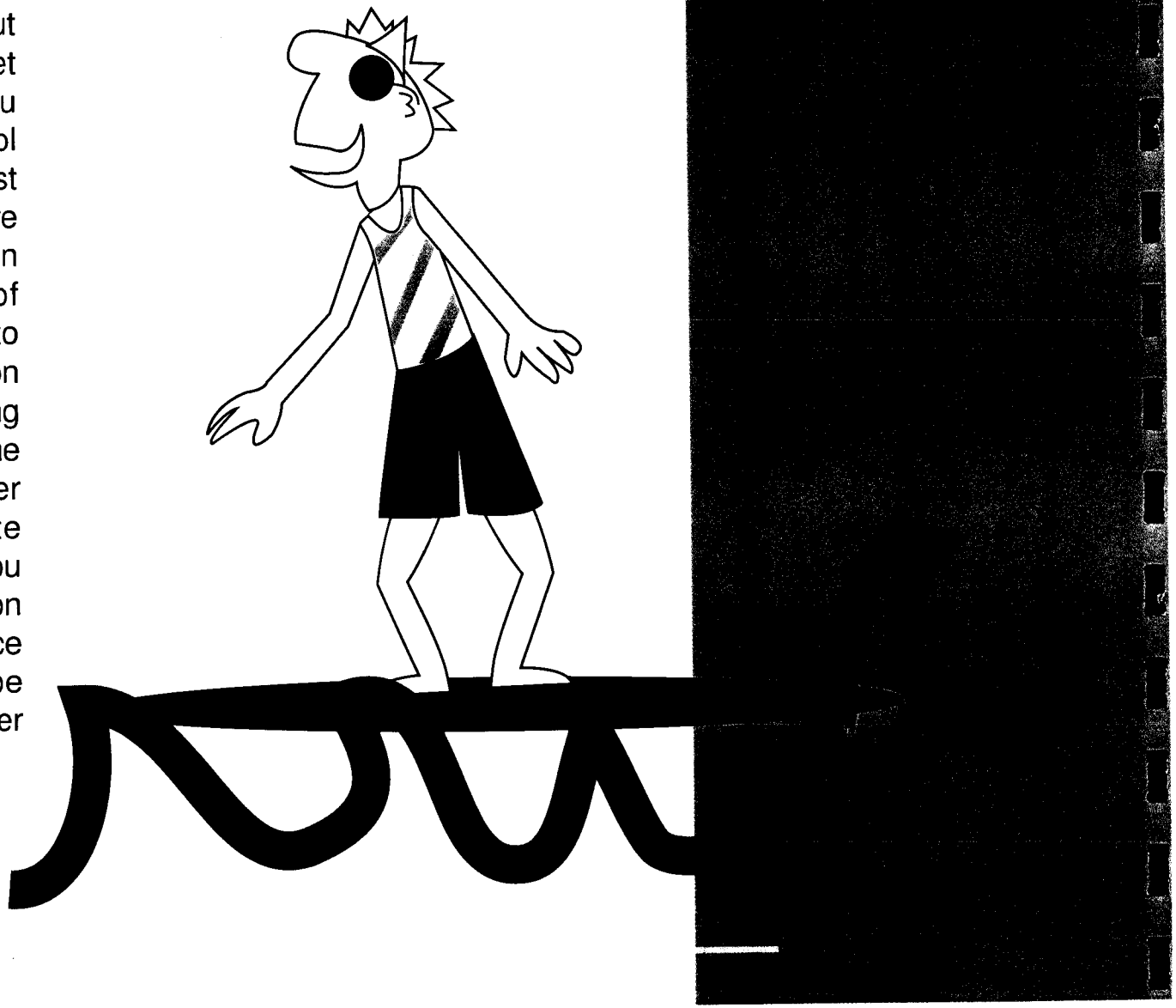
CHAPTER 3

Surfing and Search Engines

Surf's up, dude! You now know all about computers and the Internet, so let's get started with surfing the Internet. You would never jump into a swimming pool without knowing how to swim, so you must first learn how to surf the Internet before you jump into and ride the information super-highway. The Internet is full of information. There are many ways to access this information. Depending on what type of information you are looking for, some ways are better than others. The first thing to remember is that no matter how difficult it may seem to locate information you want, don't give up. If you give up, you'll never find the information you want. The Internet is a huge resource and finding your way around can be difficult, but don't give up. After this chapter you will be able to find any topic you wish, search anywhere on the Internet, and use advanced search techniques to locate that hard-to-find information. Here are some terms with which you'll need to become familiar so that you can fully understand the material within this chapter.



T

E

R

M

S

Terms

- Surfing - Surfing is the action of navigating through the Internet from site to site. Surfing can be done for many reasons. People surf the Internet for fun, to find information, and to buy products.

- Search Engine - A search engine is a piece of software on a web site that allows users to retrieve information by using a keyword search. The search engine will locate sites on the WWW that pertain to your keyword search.

- Web Crawlers - Web crawlers are a type of search engine. They conduct a search by looking at all resources found within a site and then catalog the information to provide a search that returns a large number of sites.

- Cataloging - Cataloging is the process of updating newly found resources on a web site that can be associated with related topics. Cataloging is used in web crawling search engines.

- Subject Index (Directory/Listing) - A subject index is a type of search engine that sets up information and lists the web sites in a card catalogue manner. This type of search engine helps users find what they are looking for by searching the web for a topic of interest and then searching sub-topics to web sites that contain that topic.

- URL (Uniform Resource Locator) - A URL is the address of a site on the Internet. A URL is the standard way of accessing resources on the Internet.
A URL looks like:
  -www.iowalegalaid.org
  -www.state.ia.us.com

- Search Results - Search results are the web sites that are returned from a search engine query. These results list the title of the site and usually a brief description of the information found within the site.

- Query - A query is a technique for retrieving information within a search engine. A query can be one word, "Gum," or it could be a boolean query, "Gum AND BubbleYum," or it could be a phrase, "BubbleYum Gum."

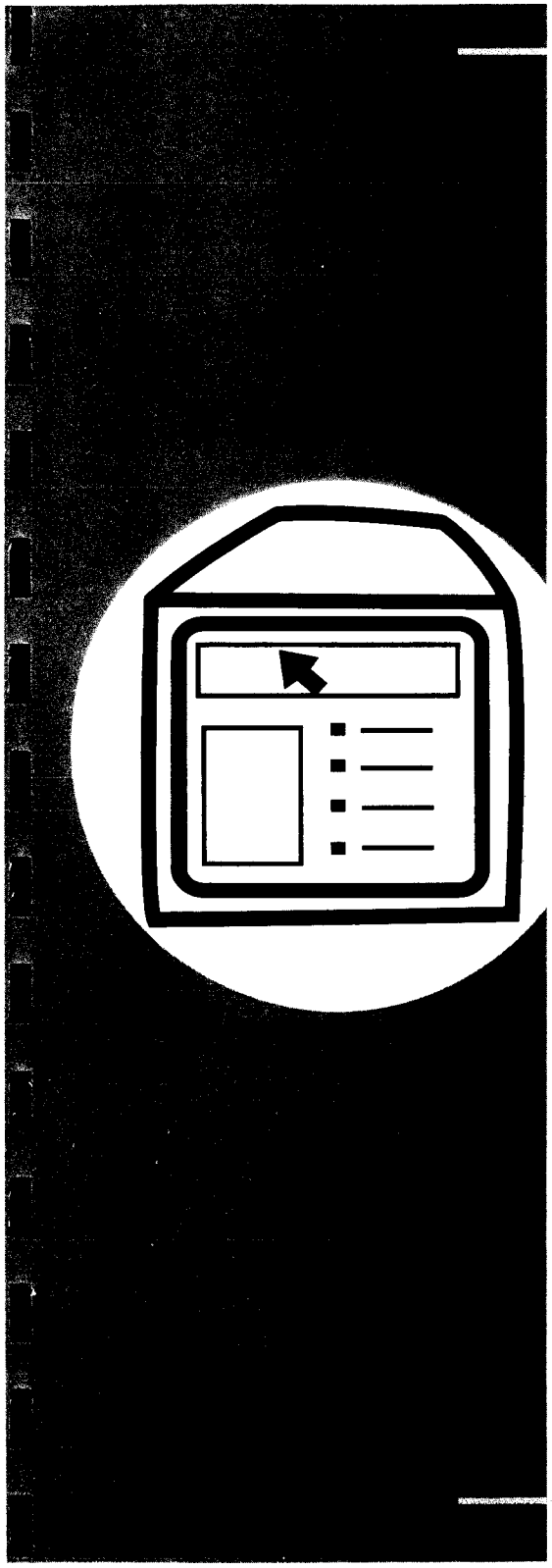
- Phrase Searching - Phrase searching is a useful way of finding information through a search engine. Typically a topic of interest will be contained within quotes. For example, to conduct a phrase search on health care issues related to Iowa, use "Iowa Health Care Issues" as your phrase.

- Boolean Searching - Boolean searching is a technique supported by most search engines that combines keywords with boolean operators. These operators are AND, OR, NOT. Boolean searching is used to help you expand or narrow your search to retrieve the most useful information.

- Nesting - Nesting is an advanced boolean search technique that combines more than one boolean operator (and, or, not) and more than two terms. An example of this type of search would be the phrase "cattle but not bulls."

Note * This information is difficult to understand and is the most complex information within the guide. Reading this material a couple of times is highly suggested and doing the exercises a couple of times is strongly encouraged. Although it may be difficult, this chapter will allow you to find anything and everything you want on the Internet. *

TERMS



Surf's Up

Now comes the fun part. You've gotten on the Internet and you don't know where to start. When you are surfing the World Wide Web, there are millions of sites that you can visit.

Your browser, Microsoft Explorer or Netscape Navigator, has several different ways of letting you access these sites. The techniques you are about to learn can be used if you know the URL of the site you wish to see. If you do not know the URL, we will discuss techniques to obtain the information you want later in the chapter.

The first way to access a specific URL is to click on the word "File" above the navigation tool bar. Once you click on this, select Open. Now you will see a dialogue box that asks you to enter or type the URL of the file you would like to open. What this is asking you to do is to type in the URL of the site you wish to visit. So, if you wish to obtain information about the Iowa Cubs, you would type in www.iowacubs.com.

The other way to access a site that you already know the address is to type it directly into the address bar on the main menu. This is the easiest way to reach a specific site. The menu bar also has a little button on the right side with a down arrow. If you click this down arrow button, you will see all of the web sites that have been typed into the address bar. This is very convenient to find addresses that you visit frequently.

If you don't know the URL of the site, you will need to use a search engine. To do this, you must first understand and pick your resources.

Picking Your Resources

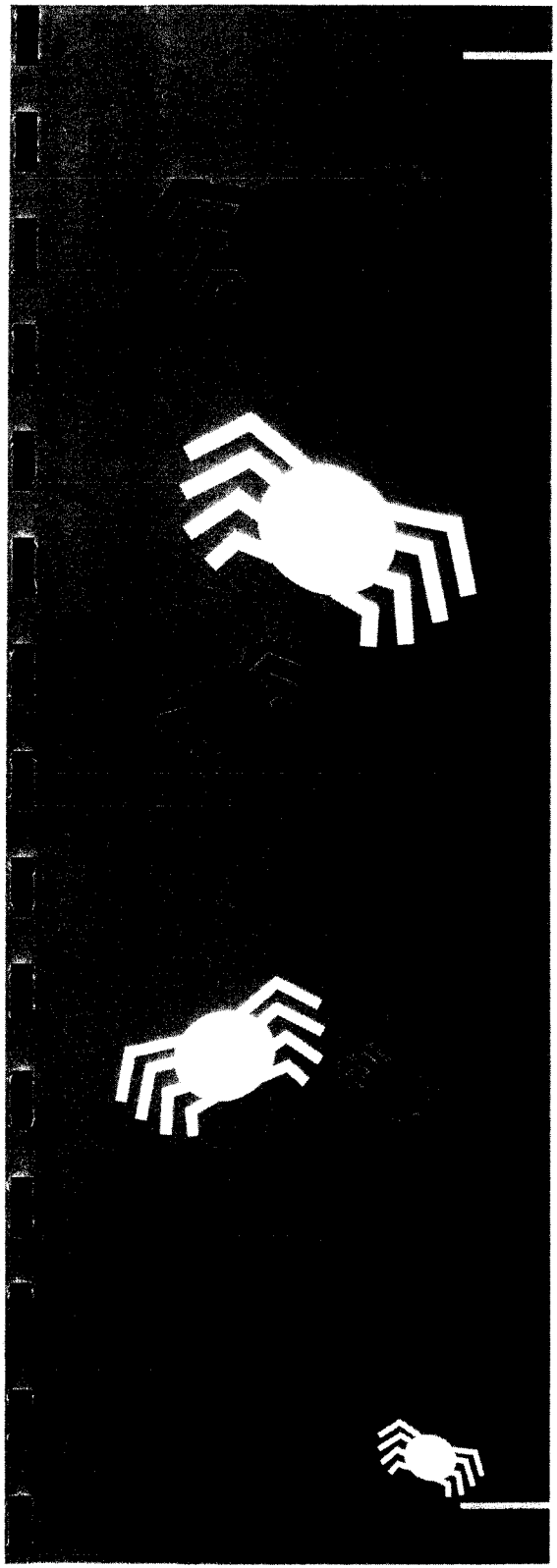
Most of the time when you are surfing the web, you won't know the URL or address of the site you wish to see. The first step to making sure you are productive on the WWW is to understand how to find information. The WWW is a large bank of resources.

Everyone has specific needs that require a specific resource. The first major resource of the WWW that you should become familiar with is a Search Engine.

Search Engines

Search Engines are a function of the WWW that allow you to locate information. Finding the exact information that you need from the WWW is a difficult task. Search engines allow you to use specific techniques to filter out the information that you don't want, and to find the information that you do want. Search Engines allow you to type in a subject and view all possible web sites that match your search. The subject that you type in is called a query. The information that is given to you is called your search results. Your search results are directly dependent on how wide or narrow you make your query. For example, you would receive a large number of search result if you were to use the query "Cars." You would receive a very small number of search results if you wanted to find information on cars but you used "Chevy 1980 Malibu Classic" as your query. Search engines are one of the most useful features of the WWW. There are two major types of search engines: Web Crawlers and Subject Indexes/Listings. Below you will find a table listing some of the search engines found on the WWW.

Site Name	Site URL	Type of Search Engine	Rating	Comments
Yahoo	www.yahoo.com	Subject Index	3 stars	Good for most information
WebCrawler	www.webcrawler.com	Web Crawler	1 star	Returns information that isn't needed
Hot Bot	www.hotbot.com	Web Crawler	1 star	Same as Web Crawler
Alta Vista	www.altavista.com	Web Crawler	4 stars	Returned the least sites and was very accurate
Dogpile	www.dogpile.com	Subject Index	5 stars	Lets you use a combo of 5 search engines!
Google	www.google.com	Web Crawler	4 stars	Very good, finds the right sites



Web Crawlers (examples)

A Web Crawler (sometimes called a spider) is an automated search engine. The way that a web crawler works is when someone submits their web site to a web crawler search engine, the search engine sends back a program telling it what type of information is found on the newly submitted site. If the program finds new resources on the submitted site, it is cataloged. This is done so that both sites can find additional links of information between them and return the best search results possible. Because both sites know the content of the other site, sometimes you will find information joined with other related information.

The advantage to this type of site is that you can find information about related topics on sites that are not dedicated to the specific subject for which you are looking. For example, if you were interested in outer space and space shuttle launches, you might surf to NASA's web page. While you are there, you might find information about airplanes that you also find interesting. Web crawlers are good information resources if you have a general topic you wish to learn about and also want to find information that goes along with your topic.

The disadvantage of web crawlers is that sites know all of the information about each other, allowing the sites to catalog too much information and make searching on these web crawlers inefficient for some specific topics.

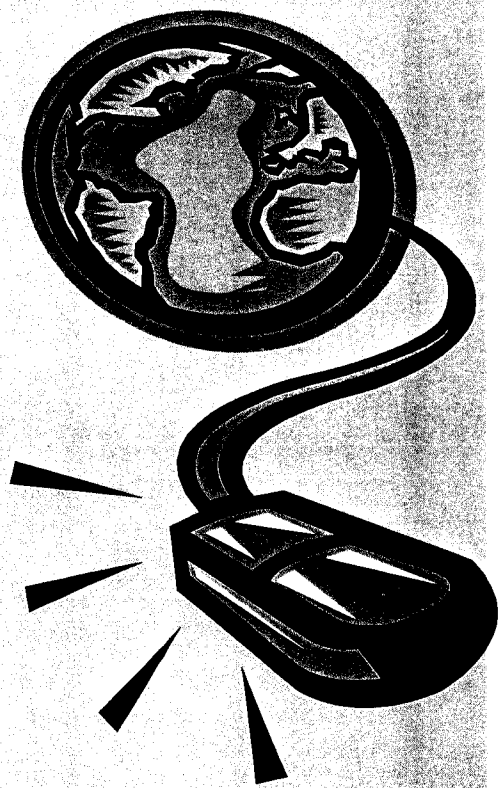
S g

Subject Indexes/Listings can be thought of as the Yellow Pages of the Internet. These search engines allow you to enter in a subject or search for the subject yourself. If you enter in the subject, the search engine will find the subject by looking at all of its directories and indexes. If you look for material yourself, you are required to narrow your search according to the directories and indexes. Subject Indexes contain only the information that is given to them. Here you will find a vast listing of resources, but if the person submitting the information about their web page doesn't submit all the information, you may not find what you want. Subject Indexes are a really good resource if you know specifically what you are looking for and if you want a wide variety of information on a topic.

Here is an example:

Let's say that your favorite National Football League team is the Oakland Raiders. Knowing that the Oakland Raiders is a National Football League team, you could use the listings provided by the search engine to find information on the Oakland Raiders. For this example we are going to use www.yahoo.com to find the information. Using the Index found on Yahoo's home page, you would start at the listing Recreation and Sports. After you clicked on this you will see more options. Next you would choose Sports. Again you would see more options. Then you would click Football (American). Upon clicking this you would see more options. Next you would click leagues. This would show you all the football leagues within the United States. Next you would click NFL, and then you would see more options. Click on the teams link, and this will take you to all of the teams in the NFL. By clicking the Oakland Raiders link, you will see all the web pages devoted to the Oakland Raiders.

PERSONAL LIST



If you were to enter "Oakland Raiders" into the search field, the search engine would do all of the work for you, but one thing that you must remember when using a Subject Index or listing resource is that you may be getting more sites than you really want.

There are some things about search engines that make them very powerful tools, and there are some bad things about search engines that make them difficult to use. Once you understand these pro's and con's, you will be able to more effectively search the web and understand how to find exactly what you want.

Pro's

Search engines provide access to a large portion of the publically available web pages on the WWW. Search engines are the single most efficient tool to help surfers find what they are looking for on the web. There is no standard on how information is kept on the web. The WWW isn't like your local library and its card catalogue system. Search engines use a variety of techniques for finding information. This is good for surfers because you can look for a specific topic on multiple search engines and find information on one search engine that you didn't find on the others.

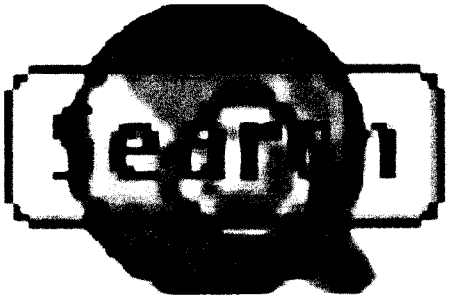
Con's

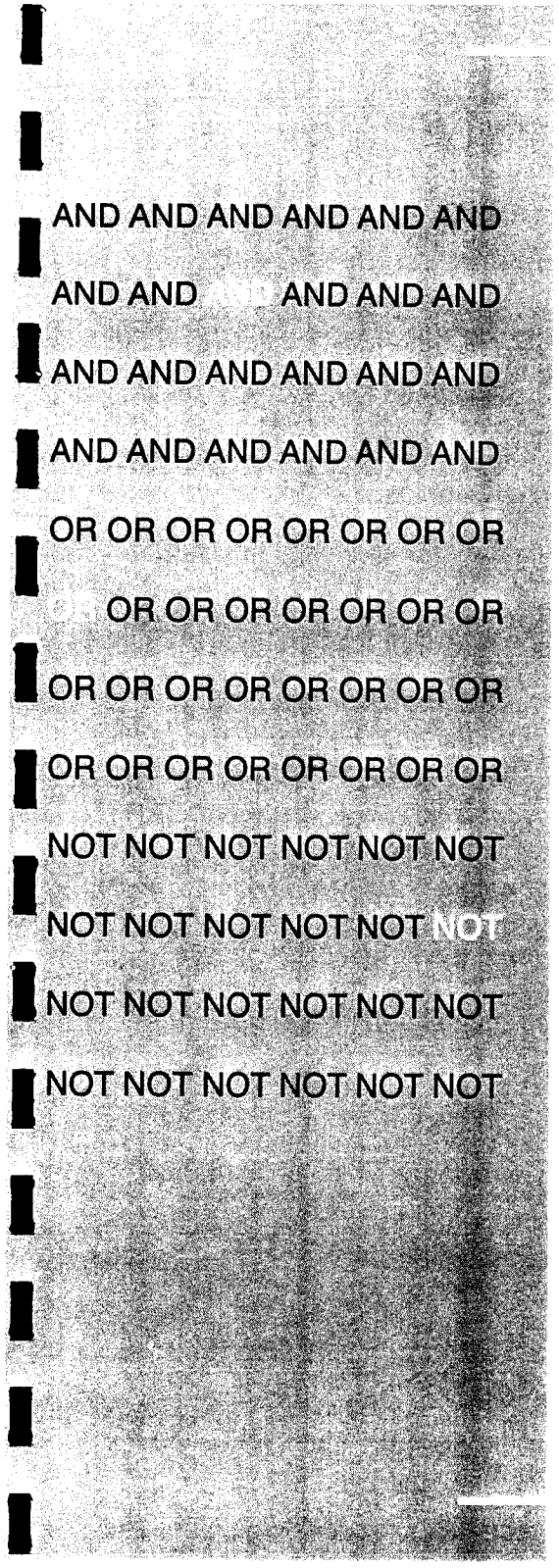
Although search engines provide a wonderful way of accessing information on the WWW, there are some negative aspects. One downside is that the more words you use to search, the more results you will receive. For example, if you were to search the phrase "Apartment for Rent," the search engine will return all documents on the web that contain any or all of those words. This increases the chances for the search engine to return literally hundreds of thousands of sites.

In order to help you find what you are looking for, we have listed some ways to conduct searches. There are many different techniques to query the data on the WWW, so let's get started and learn those techniques.

Searching can be very simple if the subject you are looking for is broad. For example, it would not be hard to find information on "Iowa." You would get back more than enough information on the state of Iowa and be able to connect to all the web sites that you wanted. The difficulty comes when you are searching for something that is very specific and requires the use of more than one word. You run into problems when you search by using more than one word because you obtain documents with all of the words or just one of the words. Most search engines and directories offer some form of advanced search, but knowing which search engine uses which type of advanced search can be hard. The two most commonly accepted advanced search techniques are phrase searching and boolean searching.

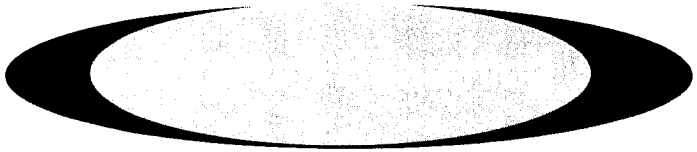
The easiest of the searching techniques is phrase searching. Phrase searching allows you to search using a combination of words. You might be doing a history assignment on Dr. Martin Luther King, Jr. and you wish to do research on him through the Internet. If you were to just type in Dr. Martin Luther King, Jr., you would retrieve sites from everywhere. Some of these sites would be about Martin Luther King, but you would get others that would talk about B.B. King, Kings from other countries, Martin Luther, and other random sites. This happens because the search engine is looking for any site that contains one or all of the words: doctor, martin, luther, king or jr. Searching like this will eventually return results that get you the information you are looking for, but it





takes time. Using phrase searching can allow you to obtain the information you want without getting so much unrelated information.

If you want to use phrase searching, you must tell the search engine that you are using phrases. The easiest way to do this is to enclose your search in quotations. So, for Dr. Martin Luther King, Jr, you would type in "Dr. Martin Luther King Jr." Because you have enclosed what you are searching for within quotes, you are telling the search engine that you wish to find documents with only "Dr. Martin Luther King, Jr." in them. This is a very useful method for searching.



Boolean searching is an advanced and very powerful way to search the WWW. The technique is named after a mathematician. Boolean searching uses a set of words called logical operators. These operators are AND, OR, and NOT. These logical keywords allow you to include information, exclude information, and separate information. These tools allow you to take full advantage of a search engine.

AND - Do a search of Term 1 AND Term 2

OR - Do a search of Term 1 OR Term 2

NOT - Do a search of Term 1 NOT Term 2

For example, if you had to do a paper on past U.S. Presidents, you could do these searches:

- Washington AND Lincoln - this would give you pages where both "Washington" and "Lincoln" appear. The boolean word AND narrows your search by selecting only those documents that have exactly the words you have listed. You can narrow your search as much as you would like by adding more and more AND's.

- Washington OR Lincoln - this would give you pages where either "Washington" or "Lincoln" appear. The boolean word OR expands your search by

retrieving documents that include all of the terms that you have listed. You can increase your search as much as you would like by adding more and more OR's. The best way to do an OR search is to start off with a very specific term and then become more general.

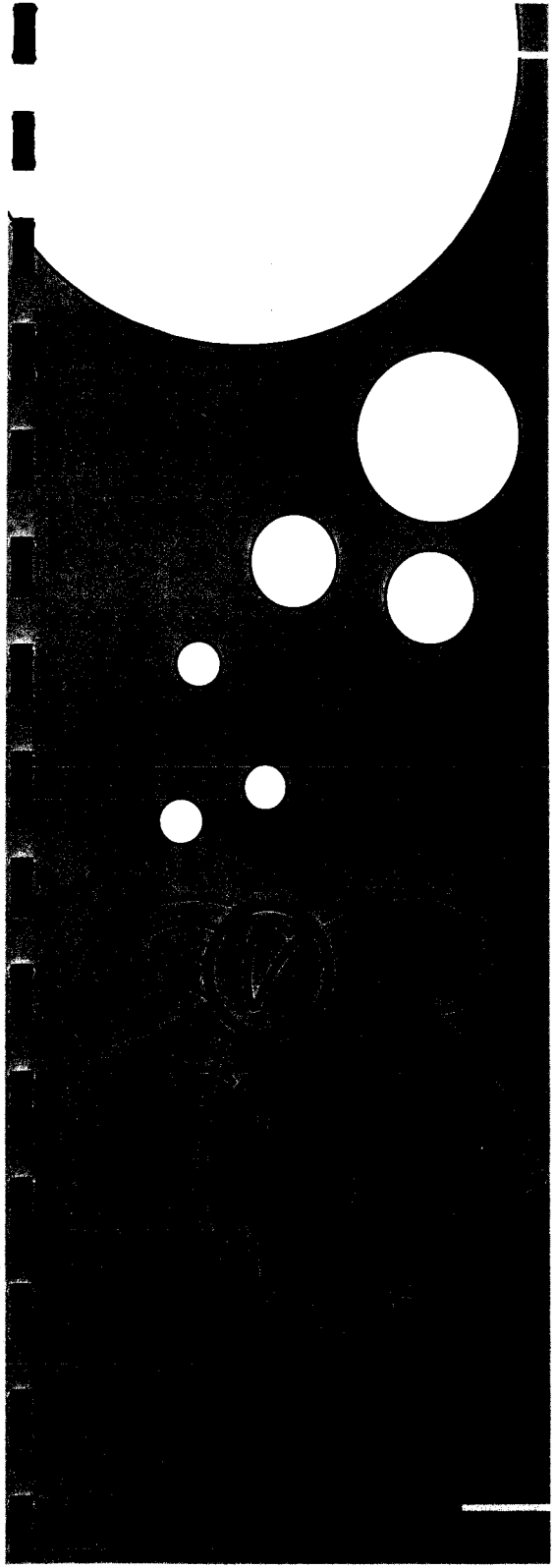
- Washington NOT Lincoln - this would give you pages where "Washington" appears and "Lincoln" does not. The boolean word NOT limits your search by selecting only those documents that contain the first term. Even if the first term appears in a document with the second term, the document will not be selected for your search.

Ne ors

Now that you know how to use the three boolean operators, let's move on to something even more powerful. You can combine boolean operators in as many combinations as you want. To do this, you need to separate the boolean keywords and group parts of your search together using parentheses. It is a good idea to use parentheses when you are using more than one operator or more than three terms that you are searching. For example:

- (Washington AND Lincoln) NOT Kennedy - this would give you pages that contain the word "Washington" and the word "Lincoln." Also, the pages from this search that contained "Kennedy" would not be retrieved. This type of nesting is a tool and can allow you to narrow your search down to exactly what you want to find.

Even though you now know how to find any subject on the WWW, there are some other tricks to help you perform the most efficient search possible. One thing to always remember is that you should start your search being very specific. If you don't find what you are looking for, you should broaden your topic. Continue doing this process until you reach a



related subject that has information about your subject. Another helpful tip is to cut out any "stop words." These words include: a, an, as, at, be, if, on, of, with.

Congratulations!!! You've just made it through the most difficult section of the Internet Guide. Give yourself a pat on the back, finish reading the chapter and then do the exercises to practice what you have learned.

Creating a Search Strategy

You have all this knowledge about how to find what you want, but you still don't know where to start. This section is designed to help you develop a strategy for finding all that information. This section is very important and pretty easy to follow. It will provide you with a simple method to find all the information you want. Here goes...

Getting started - First, before you do anything, you have to THINK about what you actually want to research or find. Are you just going to browse for some information, are you going to search a subject, or are you just trying to gather as much information about something as possible? Think about this before you start your search.

Forming your query - When you are putting together your query, remember all of the boolean operators and the use of phrase searching. Try to use at least three related words within your query. Combine your key words into phrases, and remember, you can find anything that you want. Before you search, try to think of words that would be in the body of the page and use them as keywords in addition to the topic.

Make sure you get what you want - It is very important that you conduct your search at more than one search engine. This way you know that you are getting the best sites from all search sites on the Internet. If you get a large number of responses to your search, you probably need to refine your search a little and get better results.

Test Your Knowledge

1. Supply the missing boolean operator to satisfy the following statement.
- Baseball _____ Ken Griffey Jr. (Web pages with both baseball and Ken Griffey Jr.)
- Wars _____ Civil War (Web pages about wars but without the Civil War)
- Nutrition _____ Health (Web pages about nutrition and web pages about health, but it doesn't have to be both together)
- Money _____ US (Web pages about money other than currency within the US)

2. Go to www.yahoo.com. This exercise is to show you how you can eliminate unwanted information by use of phrase searching. Type in Continental Congress. Notice the number of sites that come back and that some of them are not related to the Continental Congress. Now do the same search but put quotes around your search. "Continental Congress." Notice the difference.

3. Pick a subject that interests you. Go to www.yahoo.com. From here, try to find information about your subject using the indexes. Do not use the search engine. This exercise is to show you how indexes work. Hint: you need to first figure out in which category your subject will be located (Recreation and Sports, Computers, Education, etc.).

4. Pick a search engine from the list above and research the following topics. Record the URL's of the web sites if you wish.

- Pets that are not cats - Mother's Day - Benjamin Franklin and his inventions - The history of Iowa - Your basic Human Rights